

# Методы оптимизации в машинном обучении

Иванов П.С., Петрова А.К.

*Научный руководитель: к.т.н., доц. Сидоров В.П.*

## Аннотация

В статье рассматриваются современные методы оптимизации, применяемые в задачах машинного обучения. Особое внимание уделяется алгоритмам градиентного спуска и их модификациям. Приводятся сравнительные характеристики различных подходов к оптимизации нейронных сетей.

## Введение

Машинное обучение<sup>1</sup> стало неотъемлемой частью современных технологий. Одной из ключевых задач при построении моделей машинного обучения является оптимизация функции потерь [1].

Эффективные методы оптимизации позволяют значительно ускорить процесс обучения моделей и улучшить их качество. В последние годы появилось множество новых алгоритмов, адаптированных для работы с большими объемами данных [2].

Актуальность темы обусловлена растущим интересом к глубокому обучению и необходимостью разработки эффективных методов тренировки сложных моделей.

## 1 Основные методы оптимизации

### 1.1 Градиентный спуск

Градиентный спуск является одним из наиболее распространенных методов оптимизации в машинном обучении. Основная идея

---

<sup>1</sup>Machine learning — раздел искусственного интеллекта

заключается в итеративном обновлении параметров модели в направлении, противоположном градиенту функции потерь:

$$\theta_{t+1} = \theta_t - \eta \nabla J(\theta_t)$$

где  $\eta$  — скорость обучения,  $\nabla J(\theta_t)$  — градиент функции потерь.

**«Градиентный спуск остается фундаментальным методом оптимизации, несмотря на появление более сложных алгоритмов»** — отмечают исследователи в области машинного обучения.

## 1.2 Стохастический градиентный спуск

Стохастический градиентный спуск (SGD) является модификацией базового алгоритма, которая использует не весь набор данных для вычисления градиента, а лишь отдельные примеры или мини-батчи:

$$\theta_{t+1} = \theta_t - \eta \nabla J(\theta_t; x_i, y_i)$$

Это позволяет значительно ускорить процесс обучения на больших наборах данных [3].

## 1.3 Адаптивные методы

К адаптивным методам оптимизации относятся алгоритмы, которые автоматически настраивают скорость обучения для каждого параметра модели. Наиболее известными представителями являются:

- AdaGrad
- RMSProp
- Adam

Алгоритм Adam в настоящее время является одним из наиболее популярных методов оптимизации в глубоком обучении.

## 2 Сравнительный анализ

Проведенные эксперименты показывают, что различные методы оптимизации демонстрируют разную эффективность в зависимости от конкретной задачи и архитектуры модели.

Для простых выпуклых функций лучше всего подходит стандартный градиентный спуск, в то время как для невыпуклых задач, характерных для глубоких нейронных сетей, более эффективными оказываются адаптивные методы.

## Заключение

В статье рассмотрены основные методы оптимизации, применяемые в машинном обучении. Показано, что выбор алгоритма оптимизации существенно влияет на скорость сходимости и качество конечной модели.

Перспективным направлением исследований является разработка специализированных методов оптимизации для конкретных архитектур нейронных сетей и типов данных.

## Список литературы

- [1] *Bishop C. M.* Pattern Recognition and Machine Learning. — Springer, 2006.
- [2] *Goodfellow I., Bengio Y., Courville A.* Deep Learning. — MIT Press, 2016.
- [3] *Bottou L., Curtis F. E., Nocedal J.* Optimization Methods for Large-Scale Machine Learning // SIAM Review. — 2018. — Vol. 60. — P. 223–311.